National Cancer Institute

The Cancer Biomedical Informatics Grid® (caBIG®) Resource Guide

Developed by the caBIG®
Documentation and Training Workspace

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES

National Institutes of Health

Introduction

Welcome to the cancer Biomedical Informatics Grid® (caBIG®) Resource Guide. The caBIG® program offers a variety of resources to meet the diverse needs of audiences across the biomedical research enterprise. To help show the connection between caBIG® resources and these needs, this guide presents different scenarios that newcomers bring to caBIG®. Each scenario includes a list of the ***most highly utilized resources*** offered by caBIG® to meet the needs described in the scenario. The guide also includes summary lists of support resources to help with a range of scenarios. We look forward to connecting with you as you **connect with caBIG®**.

Scenarios

Support Resources Summaries

Learning About caBIG®

Perhaps you have heard of caBIG® as a biomedical informatics resource but are uncertain of the scope of program offerings and support, or aren't sure where to start. For example, you may have heard that you need to be caBIG® compatible or compliant to fulfill a grant requirement, or you are working on a research project that would benefit from introducing specific informatics tools to meet study goals. You need a first look to explore the range of caBIG® possibilities.

Top Recommended Resources:

- **caBIG® Essentials Training Module**: Self-guided training program that introduces important terminology and key concepts and describes the different ways of connecting with caBIG®. https://cabig.nci.nih.gov/concepts/essentials

- **caBIG® Product Representatives**: caBIG® Product Representatives can provide your organization an overview tour of specific caBIG® tools in the life sciences and clinical research domains. Send your request to cabigproductrep@nih.gov.

- **Finding caBIG® Tools**: More than 40 tools are available for adoption from caBIG®. The caBIG® program has grouped many of its most popular tools into "Tool Bundles," collections of complementary tools in the life sciences and in clinical research that are well-supported by caBIG® support resources. https://cabig.nci.nih.gov/adopt

- **caBIG® General Forum**: An electronic bulletin board for posting your own scenarios, questions, and comments. Access the forum at: https://cabig-kc.nci.nih.gov/CrossKC/

Want additional resources? Start with the caBIG® Community Website at https://cabig.nci.nih.gov/.

Assessing Readiness and Needs

After learning about tools, standards, and infrastructure available from caBIG®, you are encouraged about their potential fit with a range of research needs in the clinical, biospecimen, *in vivo* imaging, and molecular analysis domains. You now want to determine your organization's readiness and capability to connect with caBIG® and identify the critical stakeholders to involve in next steps.

Top Recommended Resources:

- **The caBIG® Deployment Self-Assessment and Goals**: This self-guided assessment helps an organization determine how caBIG® tools and infrastructure might be deployed to meet research needs. These resources will help match your existing capabilities and needs with the software and support available from caBIG®. https://cabig.nci.nih.gov/concepts

- **Getting Connected**: This section of the caBIG® Community Website includes resources specifically designed to support your connection with caBIG®, including examples for Cancer Centers and a "Talking to Your Center" series of presentations to help craft your path forward. https://cabig.nci.nih.gov/getting_connected

- **caBIG® Knowledge Centers**: Knowledge Centers (KCs) are the nexus for an expanding community employing caBIG® tools, standards, and infrastructure in specific domains. Visit the KC Website to access demonstrations, training, frequently asked questions, and other resources. https://cabig.nci.nih.gov/esn/knowledge_centers

- **caBIG® Support Service Providers**: Support Service Providers (SSPs) are independent organizations that provide client-specific caBIG® support under negotiated client-provider business arrangements. Meeting specific criteria for performance, SSPs can help you deploy caBIG® on both large and small scales. https://cabig.nci.nih.gov/esn/service_providers

Preparing for a Clinical Trial

Your organization will be initiating multiple clinical trials, both within your own site and in collaboration with other organizations. To achieve study goals, you will all need an interoperable clinical trial framework that can be used to collect, manage, and exchange patient information, with appropriate privacy and security measures in place. Specific needs include the ability to screen patients for study eligibility, record informed consent, and register and randomize participants; generate, manage, and track individual patients' study calendars; view and transfer laboratory results; enter and manage adverse events, and report to regulatory authorities; collect and manage protocol Case Report Form data; and generate clinical data reports.

Top Recommended Resources:

- **caBIG® Product Representatives**: caBIG®'s Clinical Trials Representative can provide you with an overview tour of specific caBIG® tools in the clinical trials domain. Send your request to cabigproductrep@nih.gov.

- **The caBIG® Clinical Trials Suite**: The Clinical Trials Suite is an enterprise clinical trials management solution designed primarily for use at clinical trial sites. The Suite is comprised of a collection of interoperable modules covering a broad range of key areas in clinical trials management. https://cabig.nci.nih.gov/adopt/CTCF/

- **The Clinical Trials Management Systems (CTMS) Knowledge Center**: Learn about the tools that caBIG® offers for clinical trials management, and download tools of interest with associated documentation and training. Participate in forum discussions for end users and developers, and contribute to the domain knowledge base. https://cabig-kc.nci.nih.gov/CTMS/KC/index.php/Main_Page

- **caBIG® Support Service Providers**: Support Service Providers (SSPs) are independent organizations that provide client-specific caBIG® support under negotiated client-provider business arrangements. https://cabig.nci.nih.gov/esn/service_providers

Connecting Biospecimen Resources

The Director of your large distributed organization would like to make the various biospecimen repositories held in labs across the organization interoperable, so that departments can better share resources with each other. Existing databases range greatly in complexity. Some departments welcome a new tool to help manage their biospecimens; others wish to keep their existing tool but are willing to adapt it to be interoperable with others. You need to identify the resources that caBIG® offers to help your organization achieve the goal of interoperability.

Top Recommended Resources:

- **The Tissue/Biospecimen Banking and Technology Tools Knowledge Center**: Learn about the tools that caBIG® offers for biospecimen inventory, tracking, and basic annotation; and learn how to extend existing tools to be interoperable. Participate in forum discussions for end users and developers, and contribute to the domain knowledge base. https://cabig-kc.nci.nih.gov/Biospecimen/KC/index.php/Main_Page

- **Adapting Existing Tools - caBIG® Compatibility Guidelines**: If you decide to adapt an existing biospecimen tool, the caBIG® compatibility criteria will help you ensure the tool is interoperable with others. https://cabig.nci.nih.gov/sharable/compatible

- **caBIG® Support Service Providers**: Support Service Providers (SSPs) are independent organizations that provide client-specific caBIG® support under negotiated client-provider business arrangements. Categories of services offered by licensed SSPs include Deployment of caBIG® Tools, Adaptation of Applications to be caBIG® Compatible, Help Desk Support, and Documentation and Training Services. https://cabig.nci.nih.gov/esn/service_providers

Improving the Utility of *In Vivo* Imaging

Your organization is interested in deploying imaging tools both to maximize the benefits of imaging as a tool for clinical decision support, and to better manage images across research efforts. You want a set of tools that can allow for the secure and easy sharing of images in a variety of settings, the annotation of images with clinical information, and improvements in the standardization of tumor change evaluation across researchers. Ultimately, this effort will better position the organization for large-scale collaborative research studies to demonstrate the specific benefits of *in vivo* imaging techniques on patient outcomes.

Top Recommended Resources:

- **The caBIG® Imaging Workspace**: This is the caBIG® workspace that coordinates imaging efforts for caBIG® and can provide information about the imaging tools available. https://cabig.nci.nih.gov/workspaces/Imaging

- **caBIG® Tools Search Page**: All caBIG® tools have a Tools Landing Page that provides a tool description; links to demos, documentation and training; and access to the tools files. Search for tools by domain category, such as Imaging, at https://cabig.nci.nih.gov/tools/toolsearch_view.

- **NCI-NCICB Applications Support Group**: Applications Support provides both e-mail and phone support for NCI tools, and can help direct tools-related questions to the appropriate support resources. Write to ncicb@pop.nci.nih.gov or call toll free: 888-478-4423.

Integrating and Analyzing Molecular Data

Several research laboratories in your organization wish to ease the storage and exchange of data between researchers working in the area of molecular characterization, and expand the range of potential integrative research experiments that can be conducted. In particular, you need tools that simplify microarray data storage and retrieval; enable the integration, visualization, and analysis of gene expression and sequence data; and facilitate the management of analysis results.

Top Recommended Resources:

- **caBIG® Product Representatives**: caBIG®'s Life Sciences Product Representative can provide your organization an overview tour of specific caBIG® tools in the life sciences domain, specifically related to tools to support molecular research and analysis. Send your request to cabigproductrep@nih.gov.

- **The caBIG® Life Sciences Distribution Tools Bundle**: Including a range of molecular analysis, imaging, and biospecimens tools, this tools bundle helps facilitate the discovery of the next generation of cancer diagnostics and therapeutics to realize the vision of molecular medicine. Together, they enable cancer researchers to more easily integrate, analyze, and share data from many different sources. https://cabig.nci.nih.gov/adopt/LSD/

- **The Molecular Analysis Tools Knowledge Center**: Learn about the tools that caBIG® offers for molecular characterization, including a molecular data repository, genomic analysis tools, and querying and reporting tools. Participate in forum discussions for end users and developers, and contribute to the domain knowledge base. https://cabig-kc.nci.nih.gov/Molecular/KC/index.php/Main_Page

- **caBIG® Life Cycle Demo**: This recorded demonstration uses a translational medicine example to show how several caBIG® tools and datasets can be used together to support scientific and clinical research activities. https://webmeeting.nih.gov/p29657145

Adapting a Tool for caBIG® Compatibility

Your organization has a variety of customized in-house and commercial tools that have been supporting your research effectively. Now, however, you wish to modify – or adapt – these tools to be caBIG® compatible, so they can interoperate with other tools and resources both internally and across organizations. You need to assess the steps and resources required to plan for this effort. (You'll learn about training resources in the next scenario!)

Top Recommended Resources:

- **Understanding caBIG® Compatibility**: caBIG® compatibility is about utilizing standards to ensure interoperability among research tools. Interoperability is the ability of information systems to both access and appropriately use data from a remote data resource, and it is a key goal of caBIG®. Resources on the caBIG® Website include an Adaptation Roadmap presentation, caBIG® Compatibility Guidelines, and links to the caCORE infrastructure – critical resources to support adaptation. https://cabig.nci.nih.gov/sharable/compatible

- **The Vocabulary Knowledge Center**: Access this knowledge resource to learn about tools available to effectively structure and manage shared vocabularies, a baseline requirement for compatibility. https://cabig-kc.nci.nih.gov/Vocab/KC/index.php/Main_Page

- **caCORE**: caCORE (Cancer Common Ontologic Representation Environment) provides the building blocks and tools needed to develop interoperable information management systems, enabling data sharing from the bench to the bedside and back. https://cabig.nci.nih.gov/sharable/concepts/caCORE_overview

- **caBIG® Support Service Providers**: Support Service Providers (SSPs) licensed in the Adaptation and Enhancement of caBIG®-Compatible Software Applications category may be able to support your customized adaptation efforts. https://cabig.nci.nih.gov/esn/service_providers

Accessing Training for an Adaptation Effort

Once you have identified the requirements for an adaptation effort, you may need key personnel to access training resources related to caBIG® compatibility and caCORE tooling. These training programs are designed for technical professionals with expertise in Unified Modeling Language (UML), Java development, and Service-Oriented Architecture (SOA).

Top Recommended Resources:

- **Essentials of caBIG® Compatibility Self-Guided Training**: This is a three-part self-guided training program. The first session introduces the caBIG® compatibility criteria, outlines success criteria for making a tool caBIG® compatible, and points to additional resources. The second session focuses specifically on the practical issues of semantic annotation. The third session focuses on the practical reuse of metadata. https://cabig.nci.nih.gov/sharable/compatible

- **caCORE Training Curriculum**: Teaches technical specialists how the Cancer Data Standards Repository (caDSR) fits into the caCORE Infrastructure; the role of the caCORE Infrastructure in caBIG® compatibility; and how to use tools to search for, retrieve, analyze, and curate common data elements. Tools covered include the CDE Browser, UML Model Browser, Curation Tool, Form Builder Tool, Sentinel Tool, caCORE Software Development Kit (SDK), and the Semantic Integration Workbench (SIW). https://wiki.nci.nih.gov/display/caCORE/caCORE+Training+Wiki

- **caBIG® Developer Boot Camps**: caBIG® Developer Boot Camps provide hands-on experience in using caCORE tools to create caBIG®-compatible applications, outline UML submission processes, discuss strategies for reusing metadata, and identify other resources. https://wiki.nci.nih.gov/display/caCORE/

- **caBIG® Support Service Providers**: Support Service Providers (SSPs) in the Training and Documentation Services category may be able to provide on-site or distance learning for your adaptation teams. https://cabig.nci.nih.gov/esn/service_providers

Getting on the Grid

Your organization has installed a variety of caBIG® tools to support individual and internal research efforts. In addition, existing tools have been adapted to be caBIG® compatible, using well-defined information models, controlled vocabularies, and common data elements. Now, the organization wants to take advantage of the caBIG® infrastructure backbone to "get on the Grid" to allow for inter-institutional sharing of both laboratory and clinical data resources and analytical tools.

Security is a significant concern for the organization, so connecting to any grid to enable data sharing will only happen if well-defined authentication and authorization processes are available, and if data owners are able to implement the controls needed to protect their data. The organization and its researchers must be able to define and enforce their own security policies and control access, sharing as much or as little data as the organization decides, with exactly whom they choose.

Top Recommended Resources:

- **caGrid (or "the Grid")**: caGrid is the underlying network architecture and platform that provides the basis for connectivity of caBIG® tools and data across a network of federated resources. https://cabig-kc.nci.nih.gov/CaGrid/KC/index.php/CaGrid_Overview

- **The caGrid Knowledge Center**: Access this comprehensive knowledge resource to learn about caGrid components and deployment. User tracks provide information for Principal Investigators, organizational leads, and technical specialists. https://cabig-kc.nci.nih.gov/CaGrid/KC/index.php/Main_Page

- **Data Sharing and Security Framework (DSSF)**: Tool designed to facilitate appropriate data sharing between and among organizations by addressing legal, regulatory, policy, ethical, proprietary, contractual, and socio-cultural barriers. https://cabig-kc.nci.nih.gov/DSIC/KC/index.php/Data_Sharing_and_Security_Framework

Facilitating Inter-Institutional Sharing

You are a legal or bioethics professional who works with research teams to ensure that appropriate regulations are met across the organization. Researchers have come to you with their proposal to deploy caTissue Suite, caBIG®'s biospecimen management system, to allow the sharing of tissue-related data as part of a multi-institutional collaborative research project. The researchers need to determine the steps required to be able to share their research data with others.

Top Recommended Resources:

- **The Data Sharing and Intellectual Capital (DSIC) Knowledge Center**: Provides a centralized repository of processes, model agreements, frameworks, and other resources to encourage and facilitate data sharing to advance scientific discovery, consistent with applicable legal, regulatory, ethical, and contractual requirements. https://cabig-kc.nci.nih.gov/DSIC/KC/index.php/Main_Page

- **The caBIG® Data Sharing and Intellectual Capital (DSIC) Workspace**: The mission of the DSIC Workspace is to facilitate data sharing between and among caBIG® participants by addressing a variety of barriers to data exchange. https://cabig.nci.nih.gov/working_groups/DSIC_SLWG

- **Data Sharing and Security Framework (DSSF)**: Tool designed to facilitate appropriate data sharing between and among organizations by addressing common barriers. https://cabig-kc.nci.nih.gov/DSIC/KC/index.php/Data_Sharing_and_Security_Framework

- **caGrid Knowledge Center**: Provides expertise in understanding and accessing caGrid security services when using caBIG® tools. https://cabig-kc.nci.nih.gov/CaGrid/KC/index.php/Main_Page

The DSIC Workspace welcomes new participants to contribute to the creation of new materials, and to provide feedback about data sharing and security tools and resources.

Sharing Biospecimen Data

You are a researcher with two scenarios. In the first, you are conducting a research project, and wish to assess the availability of specific types of biospecimens at other research institutions. You know that caBIG® has a range of resources to support the biomedical research enterprise, and are curious if you can identify biospecimen resources at other institutions using caBIG®. The second scenario is an extension of the first. In addition to locating others' tissues, you are also willing to post annotated data about the tissue your laboratory holds, as long as your organization has the final say in what's actually shared. If you could share this annotated data while also getting a better tool for biospecimen management, that would be ideal.

Top Recommended Resources:

- **caB2B (cancer Bench-to-Bedside)**: A caGrid software client that permits bench scientists, translational researchers, and clinicians to leverage caBIG®-compatible data and analytical services to query tissue banks at multiple cancer centers and perform in silico experiments using already-collected data sets. https://cabig.nci.nih.gov/tools/caB2B

- **caTissue Suite**: caTissue Suite is a tissue bank repository tool for biospecimen inventory, tracking, and basic annotation. caTissue permits users to track the collection, storage, quality assurance, and distribution of specimens, as well as the derivation and aliquotting of new specimens from existing ones. It provides browser-based and programmatic access to biospecimen data and manages tissue, fluid, cell, and molecular biospecimen information. https://cabig-kc.nci.nih.gov/Biospecimen/KC/index.php/CaTissue_Suite_v1.1

Want additional resources? Visit the Tissue/Biospecimen Banking and Technology Tools Knowledge Center: https://cabig-kc.nci.nih.gov/Biospecimen/KC/index.php/Main_Page

Requesting Tool Features and Reporting Bugs

You are the lead for a technical team that has implemented a number of caBIG® tools in your organization. In your development efforts, you have identified additional features for each of the tools that you would like to raise to caBIG® for possible consideration in future development efforts. You've also identified some minor bug fixes that are needed.

Top Recommended Resources:

- **caBIG® Knowledge Centers**: Each Knowledge Center manages tools forums, a bug tracker for tools that they support, and development code to enable open source development. Visit their Websites to report bugs, access code, and collaborate with other developers. Talk to the Knowledge Center(s) associated with your tools of interest about how to get your new add-on evaluated and possibly deposited. https://cabig.nci.nih.gov/esn/knowledge_centers

- **caBIG® Workspaces**: These virtual workgroups serve as the central point of collaboration, connecting stakeholder groups in a domain and setting priorities for future development. Contact workspace leads to learn about participating on regular teleconferences, where new requirements and future development efforts are identified and prioritized. https://cabig.nci.nih.gov/workspaces

- **NCI-NCICB Applications Support Group**: Applications Support provides both e-mail and phone support for NCI tools. Write to ncicb@pop.nci.nih.gov or call toll free: 888-478-4423.

Websites, Training, and caBIG® Product Representatives

The following resources may be useful at the early stages of caBIG® involvement, as you determine whether and which caBIG® resources may best fit your specific needs.

- **caBIG® Community Website**: The caBIG® Community Website includes a wide range of materials to support your learning about and engaging with the caBIG® community. https://cabig.nci.nih.gov/
- **Getting Connected**: This section of the caBIG® Community Website includes resources specifically designed to plan for a deployment of caBIG® tools and infrastructure, including examples for Cancer Centers, and "Talking to Your Center" presentations. https://cabig.nci.nih.gov/getting_connected
- **caBIG® Essentials Training Module**: This self-guided training program introduces important terminology, key concepts, and describes different ways of connecting with caBIG®. https://cabig.nci.nih.gov/concepts/essentials
- **Tool Landing Pages**: Access one-page descriptions of all the software tools offered by caBIG®. Search by intended audience, domain, maturity level, or technical complexity. https://cabig.nci.nih.gov/tools/toolsearch_view
- **caBIG Training Portal**: One-stop resource for accessing caBIG® training resources. It also includes a link to the caBIG® Learning Management System. https://cabig.nci.nih.gov/training
- **caBIG® Product Representatives**: These specialists can provide your organization an overview tour of specific caBIG® tools in the life sciences and clinical sciences domains. Write to cabigproductrep@nih.gov.
- **NCI-NCICB Applications Support Group**: Applications Support provides both e-mail and phone support for NCI tools. Write to ncicb@pop.nci.nih.gov or call toll free: 888-478-4423.

caBIG® Knowledge Centers

The caBIG® Knowledge Centers have been established at institutions with demonstrated expertise in a specific area of focus or domain of interest. They provide the following services for their domain:

- Web-based support for education, training, problem solving, and deployment needs
- Community outreach to expand interest in caBIG®
- Domain expertise for those wanting to integrate caBIG® tools and infrastructure in support of their research efforts
- Maintenance of a central repository of domain-specific tools, documentation, training, policies, and standards

Knowledge Centers include:

- **caGrid**: Provides knowledge resources for those interested in learning about, using, and contributing to caGrid.
- **Clinical Trials Management Systems**: Provides knowledge resources related to clinical trials management, and supports tools related to trial participant registration, adverse events management, patient calendar management, and other clinical data management needs.
- **Data Sharing and Intellectual Capital**: Provides centralized repository of processes, model agreements, data sharing and security frameworks, and other resources to encourage and facilitate data sharing to advance scientific discovery.
- **Molecular Analysis Tools**: Provides knowledge resources related to molecular analysis, including support for caBIG® array management and molecular analysis tools.
- **Tissue/Biospecimen Banking and Technology Tools**: Provides knowledge resources related to biospecimens management.
- **Vocabulary**: Provides knowledge resources for individuals and institutions interested in making use of or extending caBIG® tools and other vocabulary tools.

Access caBIG® Knowledge Centers:
https://cabig.nci.nih.gov/esn/knowledge_centers

caBIG® Support Service Providers

The caBIG® Support Service Providers are independent organizations that are approved by NCI as meeting specific criteria for performance. With licenses in specific service areas, they provide client-specific caBIG® support under negotiated business arrangements with the client, independent of the NCI and the caBIG® program. The caBIG® Support Service Provider licenses fall into four categories:

- **Help Desk Support**: This service category includes support for the end users and local IT administrators in institutions using caBIG® applications and tools.

- **Adaptation and Enhancement of caBIG®-Compatible Software Applications**: This service category supports caBIG®-compatible software development and includes adaptation of existing systems for caBIG® compatibility, custom enhancements to existing caBIG® tools while maintaining caBIG® compatibility, and de novo development of caBIG®-compatible applications and tools at the client's request.

- **Deployment Support for caBIG® Software Applications**: This service category includes onsite or offsite procurement and provisioning of hardware, operating systems and other software, along with installation and configuration of caBIG® software, legacy data transformation and migration. Deployment support also includes integration of multiple caBIG® tools and caBIG®-compatible applications through open application programming interfaces.

- **Documentation and Training Materials and Services**: This service category includes development of customized documentation and training materials and the provision of training on caBIG® software applications, caBIG® compatibility, and use of the caCORE infrastructure.

Access caBIG® Support Service Providers:
https://cabig.nci.nih.gov/esn/service_providers

caBIG® Workspaces

Many caBIG® activities take place within workspaces – teams that work virtually to develop technologies, white papers, guidance, and other deliverables in specific areas of interest. The workspaces serve as self-sustaining communities of stakeholders in which tools and ideas are developed, shared, and improved. Workspace participants provide data, subject matter expertise and solutions; create and evaluate standards; help prioritize activities; and help shape the workspace's direction.

- **Clinical Trial Management Systems Workspace**: Developing a comprehensive set of modular, interoperable, and standards-based tools designed to meet diverse clinical trials management needs.
- **Integrative Cancer Research Workspace**: Producing modular and interoperable tools and interfaces that provide for integration between biomedical informatics applications and data.
- ***In Vivo* Imaging Workspace**: Creating, optimizing, and validating tools and methods to extract meaning from and share imaging data.
- **Tissue Banks and Pathology Tools Workspace**: Developing tools to inventory, track, mine, and visualize biospecimens and related annotations from geographically dispersed repositories.
- **Architecture Workspace**: Developing caGrid, which supports the integration of diverse data types and interoperable analytic tools. Defines syntactic interoperability and mentors developers.
- **Vocabularies and Common Data Elements Workspace**: Evaluating and integrating systems and standards for vocabularies, common data elements and ontology content development. This workspace also defines semantic interoperability and mentors developers.
- **Data Sharing and Intellectual Capital**: Addressing issues and developing recommendations related to data sharing, patient privacy, intellectual capital, and security.
- **Documentation and Training**: Defining guidelines, processes, templates, and tools for developing consistent software documentation and training materials throughout caBIG®.

Access caBIG® Workspaces: https://cabig.nci.nih.gov/esn/workspaces/index_html

NATIONAL® CANCER INSTITUTE

NIH Publication No. 10-7518
Printed February 2010